Keywords: protein stretching, protein folding, manipulation of proteins, Go model, molecular dynamics, ubiquitin, integrin

PACS numbers: 82.37.Rs, 87.14.Ee, 87.15.-v

Stretching of Proteins in a Uniform Flow

P. Szymczak [1] and Marek Cieplak[2]

[1]*Institute of Theoretical Physics, Warsaw University,*

ul. Hoża 69, 00-681 Warsaw, Poland

[2]*Institute of Physics, Polish Academy of Sciences,*

Al. Lotników 32/46, 02-668 Warsaw, Poland

Abstract

Stretching of a protein by a fluid flow is compared to that in a force-clamp apparatus. The comparison is made within a simple topology-based dynamical model of a protein in which the effects of the flow are implemented using Langevin dynamics. We demonstrate that unfolding induced by a uniform flow shows a richer behavior than that in the force clamp. The dynamics of unfolding is found to depend strongly on the selection of the amino acid, usually one of the termini, which is anchored. These features offer potentially wider diagnostic tools to investigate structure of proteins compared to experiments based on the atomic force microscopy.

I. INTRODUCTION

The deformation of polymers in a flow has been a subject of active research for at least seventy years (see e.g.[1,2,3]). A recent renewed interest in this topic[4,5,6,7,8] arose due to development of precise experimental techniques allowing for studies of conformations at a single-molecule level. In particular, biological macromolecules such as DNA have been intensely studied in this way[9,10,11,12]. Stretching by a flow is also at the heart of the manipulation technique known as molecular combing used in genomic studies[13,14,15,16,17] and in nano-electronics[18].

In this paper, we concentrate on analysis of protein unfolding in uniform flow and compare it with unfolding in a force clamp[19], i.e. under the condition of a constant force applied to one of the termini. Theoretical studies on protein stretching in a flow are scarce[20,21] and limited to the minimalist β - barrel model. Here, we present a theoretical method to study flow induced deformations of, in principle, any protein and we illustrate it by considering ubiquitin and integrin. These two proteins were chosen because of their contrasting dynamical behavior, as established through simulations[22], when unfolding in a force clamp: ubiquitin unfolds as a rule in a single kinetic step whereas integrin – in multiple steps, i.e. with several intermediates. Additionally, we also consider synthetic α-helices, which are homopolymers, and show a very different behaviour than that seen in complex proteins.

The model we propose here is coarse grained and it involves no explicit solvent. It is an extension of the Go-like modelling[23] that we have employed in the past to study folding[24,25] and stretching at constant velocity[26,27,28]. In short, a protein is represented by a chain of C^{α} atoms that are tethered by harmonic potentials with minima at 3.8 Å. The effective self-interactions between the atoms are either purely repulsive or are minimum-endowed-contacts of the Lennard-Jones type. The parameters of these potentials reflect existence of the surrounding solvent in the sense that their minima correspond to the experimentally determined distances between the C^{α} atoms of a protein in water.

The dynamics of the protein is governed by the Langevin equation

$$m\ddot{\mathbf{r}}_\mathbf{i} = -\gamma(\dot{\mathbf{r}}_i - \mathbf{u}(\mathbf{r}_i)) + F_i^c + \Gamma \quad . \tag{1}$$

Here, $\mathbf{r_i}$ is the position of i'th aminoacid, F_i^c is the net force on it due to contact potentials, γ is the friction coefficient, and Γ is a white noise term with the dispersion of $\sqrt{2\gamma k_B T}$, where k_B is the Boltzmann constant. Finally, $\mathbf{u}(\mathbf{r}_i)$ denotes the solvent flow field.

In principle the model allows for more general description in which both γ and m are amino acid dependent and, in particular, friction is reduced for hydrophobic residues since hydrophobicity leads to a slip[29,30]. In this paper, we stay with the spirit of the traditional Go-like modelling in which all features of the protein are assumed to be contained in the native geometry and consider uniform masses and friction coefficients. Also, we neglect the effects of hydrodynamic interactions, which corresponds to the free- draining limit. This is a serious approximation, since the hydrodynamic forces between the particles contain long-range terms decaying as R^{-1} with the interparticle distance. However, since the number of residues in the considered proteins is relatively large (100- 200), the inclusion of hydrodynamic interactions (HI) into the Langevin dynamics scheme would add considerably to the numerical complexity of the problem, rendering an accurate calculation of mean unfolding times unfeasible, particularly in the small-force regime where those times are exceedingly long.

In fact, the hydrodynamic effects are very rarely taken into account in the numerical simulations of protein folding and unfolding: in the all-atom MD simulations sometimes an explicit solvent is used; but this restricts severely the largest feasible trajectory length. The coarse-grained models, in principle, would be a best starting point for the analysis of the impact of HI on the protein dynamics. However, to the best knowledge of the authors, no detailed studies of the impact of HI on protein folding and unfolding were performed, even though it was argued[31] that such an effect is expected to be non-negligible.

Keeping the above in mind, we nevertheless believe that the free-draining case may still provide useful insights on protein dynamics in a flow. This is partially confirmed by the results of the analysis of DNA streching in a uniform flow[6], where it has been observed

that the change in the extension vs flow dependence is very modest when no hydrodynamic interactions are included in the model (see also the respective discussion in[11]). A similar conclusion is drawn by Hsieh et al.[32] who compare the results of extensive simulations of the bead-spring DNA models in extensional flow to the experimental data by Perkins et al.[5]. They note that deformation-dependent HI has very little effect on the extensional flow properties of DNA molecules, whereas the rates of unraveling of single long molecule of DNA observed optically in an extensional flow can be even quantitatively predicted by beadspring models that neglect HI.

Naturally, due to their highly heterogeneous structure, proteins are much more complex than a DNA chain, thus one cannot expect those results to apply directly to the protein streching in a flow. Still, it seems that a free- draining case may provide a good starting point for understanding, at least on the qualitative level, the properties of the protein in the flow.

With the use of this simplified model, we demonstrate that flow may stretch proteins to partially unravelled stationary conformations that depend on the flow rate and on the selection of the terminus which is anchored. This is in contrast to stretching in a force clamp, in which the set of intermediate states stays the same whether we fix the C terminus and pull on the N one or do it the other way around. This difference is caused by the fact that a flow generates a non-uniform tension in the polymer. A simple explanation of this phenomenon is presented in Figure 1 for the case of a linearly positioned chain of N beads connected by bonds (in the absence of hydrodynamic interactions): the bond which is most distant from the anchor is pushed by $1/N$ of the force that is experienced by the first bond since the latter accumulates all individual pushes. It follows that, as the flow velocity is increased, the contacts near the anchored end of the protein are broken first and the protein unwinds segment by segment starting from the fixed end. Similar phenomena are observed in the experiments and simulations of polymers subject to a uniform flow (see[7] and references therein) and the corresponding shape of the partially unwound polymer was called "stem and a flower"[33,34] or "ball and a string"[35,36,37]. Thus exploration of stationary conformations corresponding to various flow rates should offer a more telling diagnostic of elastic properties of the protein than the force clamp.

We present the model in Section 2 and consider the case of the helix in Section 3. Stretching in a constant flow is analyzed in Section 4 for various ways of choosing the anchoring point and for tandem arrangement of proteins. In Section 5, we discuss an example of a non-uniform flow: elongational one. Finally, in Section 6, we consider refolding after stopping the flow.

II. THE MODEL

The effective interactions between the C^{α} amino acids are split into two classes: native and non-native. The distinction is done by checking for native overlaps of all atoms in aminoacids when represented by enlarged van der Waals spheres as proposed in reference[38]. The amino acids, i and j that do overlap in this sense are endowed with the effective Lennard-Jones potential $V_{ij} = 4\epsilon\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]$. The length parameters σ_{ij} are chosen so that the potential minima correspond, pair-by-pair, to the experimentally established native distances between the C^{α} atoms in amino acids in the pair. The repulsive interactions are described by the r_{ij}^{-12} part of the Lennard-Jones potential combined with a constant shift term that makes the potential vanish smoothly at $\sigma = 5$ Å. It should be noted that the specificity of a protein is contained in the length parameters σ_{ij}. The energy parameter, ϵ, is taken to be uniform and its effective value for titin and ubiquitin appears to be of order 900 K so the reduced temperature, $\tilde{T} = k_B T/\epsilon$ of 0.3 (k_B is the Boltzmann constant and T is temperature) should be close to the room temperature value[27,28]. All of the simulations reported here were performed at this temperature. In our stretching simulations, the anchored terminus of the protein is attached to a harmonic spring of elastic constant k=0.06 $\epsilon/\text{Å}^2$.

As explained in the Introduction, thermostating and mimicking some effects of the solvent are provided by the Langevin dynamics, Eq. 1. The friction coefficient γ is taken to be equal to $2m/\tau$ where $\tau = \sqrt{m\sigma^2/\epsilon} \approx 3ps$ is the characteristic time scale of oscillations in the Lennard-Jones well. The selected value of γ corresponds to a situation in which the inertial effects are small[24] but the damping action is not yet as strong as in water.

The equations of motion are solved by a fifth order predictor-corrector scheme. In the course of stretching, the native contacts are being ruptured. A contact between amino acids i and j is said to be ruptured if the corresponding distance r_{ij} becomes larger than $1.5\sigma_{ij}$ (close to the inflection point of the Lennard-Jones potential) for the last time. When studying folding, we consider establishing contacts starting from an unfolded state. A contact is said to be established when the corresponding value of r_{ij} crosses the threshold value for the first time. Folding is considered to be achieved when *all* contacts are established simultaneously.

When simulating the force clamp, the force F is applied to the spring that pulls one of the termini (the choice of the terminus is irrelevant in this case). In the case of the flow, we discuss the results in terms of the net hydrodynamic force, $\mathbf{F} = \gamma \sum_{i=1}^{N} \mathbf{u}(\mathbf{r}_i)$, that is experienced at the anchor point. For uniform flows, $F = N\gamma u$. The dimensionless force, $F\sigma/\epsilon$, will be denoted by $\tilde{F}$. The conformations will be characterized by the end-to-end distance L.

The relative strength of convective and diffusive effects in the dynamics of a protein is given by the Peclet number

$$Pe = \frac{UR_g}{D}$$

where U is the characteristic flow magnitude, R_g - radius of gyration and D - the diffusion coefficient of the protein. Numerically, one may estimate D by the analysis of the mean square displacement of the protein as a function of time. For example, for ubiquitin, the calculations give $D \approx 0.2\sigma^2/\tau$, whereas $R_g = 2.3\sigma$. The flow rates used in the simulations lie in the range $U = 0.02 - 0.07\sigma/\tau$, what gives Pe $\approx 0.2 - 0.7$.

Since the Peclet number is dimensionless, it can be used to relate the simulation to the experimental setup. Namely, as reported in[39], the diffusion coefficient for ubiquitin is $D \approx 1.7 \cdot 10^{-6}$ cm^2/s, whereas $R_g = 1.15 \cdot 10^{-7} cm$. Thus the above mentioned Peclet number range corresponds to the flow rates of $4 - 13$ cm/s. Such speeds are about three orders of magnitude faster than those needed to unravel DNA molecules. This is because proteins contain larger clusters of bonds that need to be ruptured simulateneously and are also smaller in size. The above comparison may also be used to relate the numerical

time scale τ to the experimental time scales. Namely, from the fact that $U = 0.02\sigma/\tau$ corresponds to 4 cm/s and $\sigma = 5$ $\mathring{A}$ one concludes that τ corresponds to approximately 0.25 ns of the 'real' time. This time scale is of the same order of magnitude as the one that Veitshans et al.[40] arrived at (3 ns) by using an entirely different argument.

III. FLOW-INDUCED STRETCHING OF HOMOPOLYMERS

A synthetic helix provides an example of a homopolymer since none of its parts, except at the termini, is distinguished. The dependence of the end-to-end distance in the stretched helix on the total stretching force is shown in Figure 2. The fractional extension is seen to be a smooth function of the applied force without any stationary or quasistationary stages. The steady-state conformations corresponding to different values of F clearly show the "stem and a flower" phenomenon – the helix unwinds from the fixed end and the unwound length depends on the net hydrodynamic force, i.e. on the flow rate. It is also seen that the dependence of the fractional change in L on F does not change with the total number of residues in the helix. This finding is consistent with a similar and well-established result[7,33,34] pertaining to homopolymers in the free-draining limit.

The other part of Figure 2 shows the dependence of the mean unfolding time of the helix on the net hydrodynamic force. For the purpose of making this figure, we consider the helix to be unfolded when its total length exceeds 90% of the maximum extension length of $(N-1) \times 3.8\mathring{A}$. It is seen that in the small-force regime,the unfolding time exponentially decreases as a function of the force. For larger forces the dependence of the unfolding time on the force becomes much weaker. An analogous phenomenon is observed in the simulations of stretching of proteins in the force clamp[22].

IV. FLOW-INDUCED STRETCHING OF PROTEINS

Figures 3 and 4 show the end-to-end distance versus time for integrin unfolding in a uniform flow. In Figure 3, terminus C is anchored, whereas in Figure 4 it is terminus N that is anchored. Several trajectories corresponding to different values of the total hydrodynamic force are shown. Since the tension is strongest near the anchoring point, the

protein unfolds from the fixed end towards the free one. In contrast to homopolymer, here the unfolding pathway traverses through a number of intermediate states, corresponding to the unzipping of subsequent structures from the bulk of the protein. We observe that if the flow rate is sufficiently low, the protein may remain trapped in one of these states for the duration of the simulation. In contrast to the simple helix, complex proteins are cross-linked and inhomogeneous and yield to inhomogeneous tension in a way which is specific to the stretching protocol. Thus the steady-state conformation in which the protein is found after a long time depends not only on the value of the force but also on the choice of the terminus. In particular, as it is seen in the Figures, the set of intermediates is much richer for the case of fixed C terminus than vice versa. Also, in the former case the full unwinding of integrin chain requires a smaller net force. This suggests that the strongest bonds in the native structure of integrin are located nearer to the C terminus.

The differences between unfolding with different termini fixed are further highlighted by analysis of the so called unfolding scenarios[24], in which one plots an average time when a given contact is broken against the contact order, i.e. against the sequential distance, $|j-i|$, between the amino acids that form a native contact. Figure 5 compares the unfolding scenarios for different anchorings of the integrin chain. Again, it is seen that anchoring at the C terminus gives rise to a much richer unfolding dynamics, including several intermediates, than the anchoring at the other terminus. Existence of these differences may offer an interesting way of the experimental probing the structure of a protein by analysis of unfolding trajectories with different anchoring points.

For comparison, Figure 6 shows the unfolding trajectories for the integrin pulled by a force applied at the terminus only, as in the force-clamp apparatus. All the intermediates present here are also seen in uniform flow experiments, particularly those with C terminus fixed (*cf.* Figure 3), but vice versa is not true. Thus uniform flow unfolding appears to be richer in intermediate conformations than simple pulling.

It is instructive to perform a similar analysis for another protein, ubiquitin. Ubiquitin behaves very differently from integrin when stretched in a force-clamp[22], since it unfolds usually in a single kinetic step whereas integrin unfolding involves several intermediates

as it is illustrated in Figure 5. However, as it is seen in Figures 7 and 8, in a uniform flow ubiquitin does display several intermediate steps in unfolding, which confirms that the unfolding in a flow shows a richer behavior than stretching in a force clamp. Additionally, one again observes a difference in behavior between C and N anchoring. This time, it is N anchoring which shows a larger number of intermediates and requires a smaller force for the full unfolding. Thus, in the case of ubiquitin, the strongest bonds are located in the neighborhood of N terminus.

Finally, Figure 9 summarizes results on statistically averaged (over 50 unfolding trajectories) flow-induced and force-clamp-induced processes of unfolding in ubiquitin (top panel) and integrin (bottom panel). The figure shows the dependence of the logarithm of the median unfolding time on the total force. Just as in the case of helix, we consider a protein to be unfolded when its total length exceeds 90% of the maximum extension length of $(N-1) \times 3.8\mathring{A}$. Again, for the uniform flow, one should note the lack of symmetry between the anchoring by the N terminus and by the C terminus. One should also observe that determination of which choice offers more resistance to unravelling is protein dependent. However, force-clamp stretching generally requires a smaller force since in the flow-induced case the segments which are near the free end are exposed to relatively small unravelling tensions and thus they unfold only partially.

V. STRETCHING OF POLYPROTEIN IN A FLOW

The experiments on protein unfolding in a force clamp are usually performed with polyprotein chains consisting of several repeats of a given protein. For example in the studies of Fernandez group[41,42], polyubiquitin chains of 2-9 linked ubiquitin domains were investigated. When a constant force is applied to the terminus of such a system, the domains unfold in a staircase-like manner, with each step corresponding to the unwinding of a single domain. Serial unwinding of polyubiquitin is also observed in molecular dynamics simulations[22]. It is observed that selection of the domain to be unravelled the first is fluctuations-driven and thus random in nature.

In a uniform flow, the situation is different, as it is seen in Figure 10 which presents the unfolding pathways for two-ubiquitin for various flow rates. First, because of the nonuniform tension along the chain, the unwinding always begins with the domain closest to the the anchoring point. Also, as it is seen in Figure 10, if the flow rate is high enough the unfolding is not serial - one of the ubiquitin domains unfolds together with a considerable piece of the other domain. For smaller flow rates, one of the intermediates corresponds to the situation when one of the domains is fully unfolded while the other is not. However, now it is just one of the many intermediate states of 2-ubiquitin and not the unique intermediate conformation, as it is the case for the force-clamp stretching.

VI. NON-TERMINAL ATTACHMENT

Since unfolding in a uniform flow depends considerably on the choice of the anchoring terminus, it is worth exploring other possibilities of anchoring. Figure 11 corresponds to an unfolding trajectory of integrin chain that is tethered at lysine 148. Here we monitor the end-to-end lengths, L_1 and L_2, of two segments (1-148) and (148-184) respectively, at the net hydrodynamic force of $\tilde{F} = 4$. We observe that, initially, both segments get streched side-by-side and at the same rate. However, as discussed in the Introduction, the tension in a longer segment is higher which leads to a rapid rupture of the inter-segmental contacts. From this time on, they evolve independently - the longer chain unwinds quickly whereas the shorter one snaps back and folds into a stationary conformation as shown in Figure 11.

VII. ELONGATIONAL FLOW

Finally, we observe that the abundance of intermediate states seen in a uniform flow is not necessarily present for other kinds of flows. As an illustration, we have carried out simulations of integrin unfolding in an elongational flow described by:

$$u_x = g(x - x_0), \quad u_y = -\frac{1}{2}g(y - y_0), \quad u_z = -\frac{1}{2}g(z - z_0) \quad . \tag{2}$$

Here, (x_0, y_0, x_0) corresponds to a location of the stagnation point for the flow (*cf.* Figure 12) and we take it to coincide with the C terminus of a protein. In over 500 unfolding

trajectories, obtained for a broad range of the elongational rate, g, we have not observed even a single stationary intermediate state, whereas there are at least four intermediate states when flow is uniform and the anchoring is applied to the C terminus. Figure 13 shows an example of the unfolding pathway in this case together with the dependence of the mean unfolding time on g.

The lack of intermediate states in elongational flow, in contrast to the uniform flow, has been already noticed by Lemak et al.[20,21]. However, they attributed the absence to the fact that "in an elongational flow every monomer experiences a force that is high enough to delocalize it from a bonding site". In our opinion, the physical mechanism here is, in fact, different. We note that, in contrast to the case of the uniform flow, the total hydrodynamic force acting on a protein chain in an elongational flow depends on the actual total length of the chain. This is so because the farther from (x_0, y_0, z_0) the pulled terminus is, the faster flow it experiences. Thus as the protein unravels even just a bit the total hydrodynamic force also increases correspondingly. This results in a positive feedback mechanism that leads to a rapid rupture until the protein is unravelled fully.

VIII. REFOLDING AFTER STOPPING THE FLOW

If the flow is stopped suddenly, the protein chain folds again. The analysis of the folding trajectories shows a considerable number of misfolded stationary conformations arising in these processes (typically about 10-20% trajectories, depending on the initial extension of the protein). Typically, the misfolded conformations lie relatively close to the native state, with a RMSD of $1 - 10\mathring{A}$ but the escape time from the misfolded conformation is often longer than the time needed to reach the misfolded state. A typical trajectory with a misfolding event is presented in Figure 14. In this trajectory, the protein gets into a conformation with essentially the same end-to-end length as in the native state without establishing about 10% of the native contacts. The lower panel of Figure 14 shows the distribution of the folding times for ubiquitin in which the initial state has been obtained by a constant flow unfolding with a total hydrodynamic force of $\tilde{F} = 5$. In this case, almost 20% of all trajectories lead to misfolding. The trajectory in the upper panel of Figure 14

corresponds to a relatively short-lived misfolded state. In most situations corresponding to misfolding, the native conformation is not reached within the duration of the simulation. Such trajectories have not been included in the histogram shown in Figure 14.

IX. SUMMARY

In summary, stretching in a uniform flow provides a promising tool for probing the conformational landscape of proteins. In contrast to homopolymers, the steady-state conformations of proteins corresponding to various flow rates form a discrete set. Unfolding usually involves several kinetic transitions between subsequent intermediates and has a richer dynamics than that in the force-clamp case. Moreover, the unfolding pathways depend on the selection of the point of anchor. Thus, making various selections provides additional information about the structure of proteins. Harnessing this information may be facilitated experimentally by attaching the free end of the protein to a fluorescent quantum dot[43]. A similar technique has been succesfully used, e.g., in tracking of the myosin molecular motor[44].

M.C. appreciates stimulating discussions with Olle Inganas. This work was supported by Ministry of Science in Poland (Grant No. 2P03B-03225) and by the European program IP NaPa through Warsaw University of Technology.

---

[1] W. Kuhn, Kolloid Z. **68**, 2 (1934)

[2] H. A. Kramers HA, J. Chem. Phys. **14**, 415-424 (1946).

[3] P. G. de Gennes, J. Chem. Phys **60**, 5030-5042 (1974).

[4] D. E. Smith and S. Chu, Science **281**, 1335-38 (1998).

[5] T. T. Perkins, D. E. Smith, S. Chu, Science **276**, 2016-21 (1997).

[6] T. T. Perkins, D. E. Smith, R. G. Larson, and S. Chu, Science **268** 83-87 (1995).

[7] R. Rzehak, W. Kromen, T. Kawakatsu and W. Zimmermann, European Physical Journal E **2**, 3-30 (2000).

[8] M. Cheon, I. Chang, J. Koplik, and J. R. Banavar, Europhys. Lett. **58** 215-221 (2002).

[9] R. G. Larson, T. T. Perkins, D. E. Smith, and S. Chu, Phys. Rev. E **55** 1794-1797 1997.

[10] S. S. Abramchuk, A. R. Khokhlov, T. Iwataki, H. Oana and K. Yoshikawa Europhys. Lett., **55**, 294-300 (2001).

[11] E. S. G. Shaqfeq, J. Non-Newtonian Fluid Mech., **130**, 1-28 (2005).

[12] T. Heim, S. Preuss, B. Gerstmayer, A. Bosio, and R. Blossey, J. Phys.: Cond. Matter **17**, S703-S716 (2005).

[13] T. R. Strick, M.-N. Dessinges, G. Charvin, J. F. Allemand, D. Bensimon, and V. Croquette, Rep. Prog. Phys. **66**, 1-45 (2003).

[14] X. Michalet, R. Ekong, F. Fougerousse, S. Rousseaux, C. Schurra, N. Hornigold, M. van Slengtenhorst, J. Wolfe, S. Povey, J. S. Beckmann, A. Bensimon, Science **277**, 1518-1523 (1997).

[15] D. Bensimon, A. J. Simon, V. Croquette, and A. Bensimon, Phys. Rev. Lett. **74**, 4754-4757 (1995).

[16] O. B. Bakajin, T. A. J. Duke, C. F. Chou, S. S. Chan, R. H. Austin, and E. C. Cox, Phys. Rev. Lett. **80**, 2737-2740 (1998).

[17] E. Y. Chan, R. Gilmanshin, R. H. Austin, E. D. Carstea, M. Fuchs, L. Gleich, N. M. Gonclaves, R. A. Haeusler, J. W. Larson, A. M. Maletta, M. P. Masera, T. G. Thompson, P. S. Wellma,, G. G. Wong, and G. R. Yantz, Bioph. J. **83**, 302A-302A, Part 2 Suppl. (2003).

[18] P. Bjork, A. Herland, I. G. Scheblykin, and O. Inganas, Nano Lett. **5**, 1948-1953 (2005).

[19] F. Oberhauser, P. K. Hansma, M. Carrion-Vazquez, and J. M. Fernandez, *Proc. Natl. Acad. Sci. (USA)* **98**, 468 (2001).

[20] A. Lemak, J. R. Lepock, and J. Z. Y. Chen, Proteins: Structure, Function and Genetics **51**, 224-235 (2003).

[21] A. Lemak, J. R. Lepock, J. Z. Y. Chen, Phys. Rev. E **67**, 031910 (2003).

[22] P. Szymczak and M. Cieplak J. Phys.: Condens. Matter **18**, L21-L28, (2006).

[23] H. Abe and N. Go, *Biopolymers* **20** 1013 (1981).

[24] T. X. Hoang and M. Cieplak, J. Chem. Phys. **113**, 8319-8328 (2000).

[25] M. Cieplak and T. X. Hoang, *Biophysical J.* **84** 475 (2003).

[26] M. Cieplak, T. X. Hoang and M. O. Robbins, *Proteins: Struct. Funct. Bio.* **56** 285 (2004).

[27] M. Cieplak, A. Pastore and T. X. Hoang, *J. Chem. Phys.* **122** 054906 (2004).

[28] M. Cieplak and P. E. Marszalek, J. Chem. Phys. **123**, 194903 (2005).

[29] M. Cieplak, J. Koplik, J. R. Banavar, *Phys. Rev. Lett.* **86**, 803 (2001).

[30] M. Cieplak, J. Koplik, J. R. Banavar, Phys. Rev. Lett. **96** 114502 (2006).

[31] H. Tanaka J. Phys. Cond. Matt., bf 17 S2795-S2803 (2005).

[32] C-C. Hsieh, L. Li, R.G. Larson, J. Non-Newtonian Fluid Mech., **113**, 147-191 (2003).

[33] F. Brochard-Wyart, Europhys. Lett. **23**, 105-111 (1993).

[34] F. Brochard-Wyart, Europhys. Lett. **30**, 387-392 (1995).

[35] B. J. Haupt, T. J. Senden, and A. M. Sevick, Langmuir **18** 2174-2182 (2002).

[36] R. G. Maurice and C. C. Matthai, Phys. Rev. E **60** 3165-3169 (1999).

[37] M. Cieplak, T. X. Hoang, and M. O Robbins, Phys. Rev. E **70** 011917 (2004)

[38] J. Tsai, R. Taylor, C. Chothia, and M. Gerstein, *J. Mol. Biol.* **290** 253 (1999).

[39] A. J. Dingley, J. P. Mackay, G. L. Shaw, B. D. Hambly and G. F. King *J. Bio. NMR*, **10** 1 (1997).

[40] T. Veitshans, D. Klimov, and D. Thirumalai. *Folding and Design*, **2** 1 (1997).

[41] J. M. Fernandez and H. Li, *Science* **303** 1674 (2004).

[42] M. Schlierf, H. Li and J. M. Fernandez *Proc. Natl. Acad. Sci. (USA)* **101**, 7299, (2004)

[43] A.R. Bausch and D.A. Weitz, *J. Nanoparticle Res.*, **4**, 477, (2002).

[44] D. M. Warshaw, G. G. Kennedy, S. S. Work, E. B. Krementsova, S. Beck, and K. M. Trybus, *Biophys J.* **88**, L30, (2005).

FIGURE CAPTIONS

Fig. 1. Schematic representation of stretching a polymer by fluid flow. The bead denoted by 1 is free whereas the bead denoted by N is attached to a spring. The other end of the spring is fixed. A uniform flow is directed from the right to the left. The tension along the chain increases linearly from the free end towards the anchored end.

Fig. 2. The top-left panel shows the dependence of the fractional extension of the helix in the steady state on the total hydrodynamic force for synthetic helices with $N = 48$ (asterisks) and $N = 24$ residues (squares). The snapshots at the bottom show examples of the stationary states for forces indicated. The top-right panel plots the logarithm of the median unfolding time against the total hydrodynamic force for the N=24 helix. For the purpose of making this figure, we consider a helix to be unfolded when its total length exceeds 90% of the maximum extension length of $(N-1) \times 3.8\mathring{A}$.

Fig. 3. Examples of the time evolution of the end-to-end distance in unfolding of integrin in a flow for the total hydrodynamic forces as indicated. The C-terminal is fixed and the conformations corresponding to the plateau regimes of the unfolding pathways are shown on the right.

Fig. 4. Similar to Figure 3 but for the situation in which the N terminus is fixed.

Fig. 5. The scenarios of unfolding of integrin in a uniform flow. The top and bottom panels corresponds to anchoring of the C and N termini respectively. The values of the total stretching forces are indicated.

Fig. 6. Similar to Figure 3 but for unfolding induced by applying a constant force, as indicated, in a force clamp.

Fig. 7. Similar to Figure 3 but for ubiquitin.

Fig. 8. Similar to Figure 4 but for ubiquitin.

Fig. 9. The dependence of the logarithm of the median unfolding time on the force. The top panel is for ubiquitin and the bottom panel for integrin. The solid data points and solid lines (the latter are guides to the eye) correspond to unfolding in a flow. The

choice of the anchored terminus is indicated next to the lines. The open symbols and dotted lines correspond to stretching in a force clamp.

Fig. 10. Unfolding pathways of two-ubiquitin for various total forces as indicated. An unfolding trajectory for two-ubiquitin in a force-clamp (the dashed line) is given for comparison. The full extension in force clamp is longer than in the flow induced case since in the latter case the segments which are most distant from the anchor experience too small force to unravel. The inset shows the scenarios of unravelling events. The contacts in the domain that is closer to the anchor are marked by black squares and those in the more distant domain by open squares. The snapshots on the right correspond to the stationary states at those values of L (approximately) at which the snapshots are plotted. The values of dimensionless forces used in the simulations ($\tilde{F}$) are indicated.

Fig. 11. The unfolding pathway of integrin anchored at $Lys148$ in a uniform flow. The end-to-end length of the segments (1-148) and (148-184) is plotted as a function of time, with the corresponding protein conformations shown. The inset shows a protein conformation just before the contacts between the segments are broken.

Fig. 12. A schemtatic view of elongational flow with a stagnation point (x_0, y_0) in the center of the graph

Fig. 13. Integrin: the unfolding pathway in an elongational flow and the median unfolding time as a function of elongational rate $\tilde{g}$ defined as $\tilde{g} = N\gamma g L_m$, where L_m is the maximum extension length of a protein, $L_m = (N-1) \times 3.8\mathring{A}$. Note that below $g \approx 1.6$ the unfolding process is essentially arrested.

Fig. 14. The top panel shows a typical folding trajectory with a misfold for integrin. The initial state was obtained by a constant flow unfolding with a total hydrodynamic force of $\tilde{F} = 5$. The time is measured from an instant at which the force is suddenly reduced to zero (i.e. the flow is stopped). Throughout the process, the C terminus is held anchored. The graph shows both the end-to-end distance (L) and the fraction of native contacts (Q). Note that the final escape from the misfolded state is only seen in the $Q(t)$ graph. The bottom panel shows the distribution of folding times in this case. The fit is to a log-normal distribution $\frac{1}{\sqrt{2\pi}\sigma(t-t_0)} \exp\left(-\frac{ln^2(\frac{t-t_0}{m})}{2\sigma^2}\right)$. with $t_0 = 7675\tau$, $\sigma = 0.4$,

and $m = 3065\tau$. The histogram is based on 300 trajectories all starting from the same stretched conformation. Trajectories with long-lasting misfolded conformations (corresponding to the unfolding times longer than 18000τ) have not been taken into account in the histogram. They make up about 20% of all trajectories.

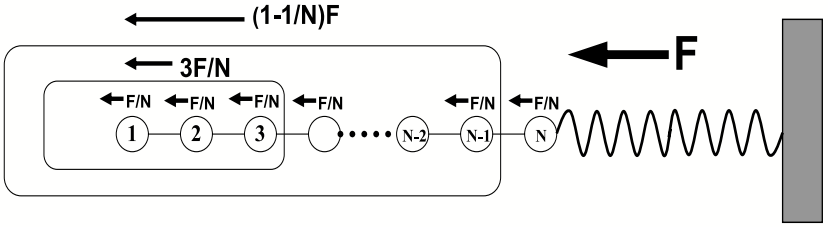

FIG. 1:

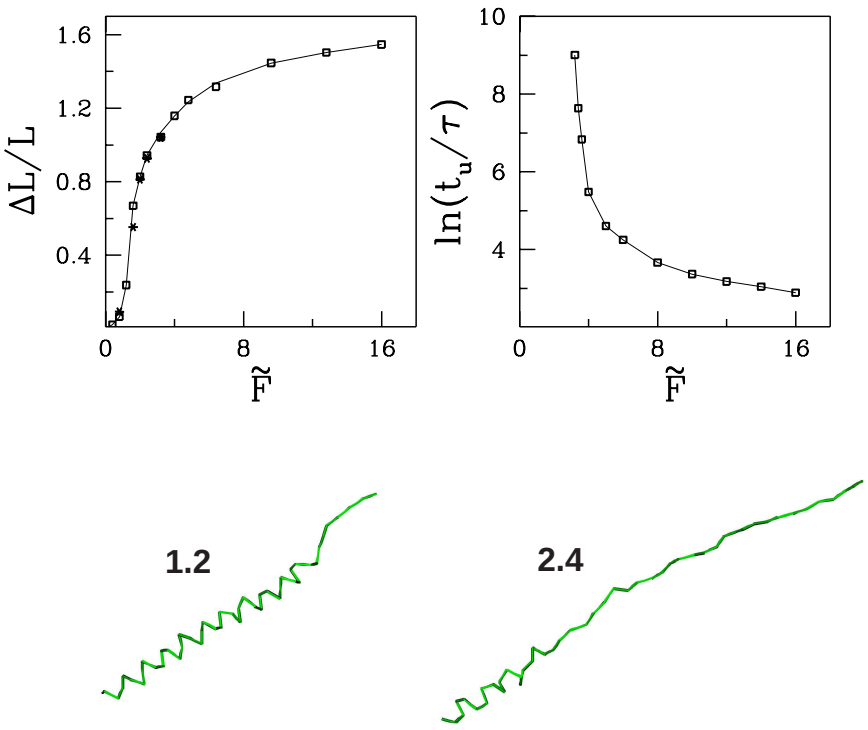

FIG. 2:

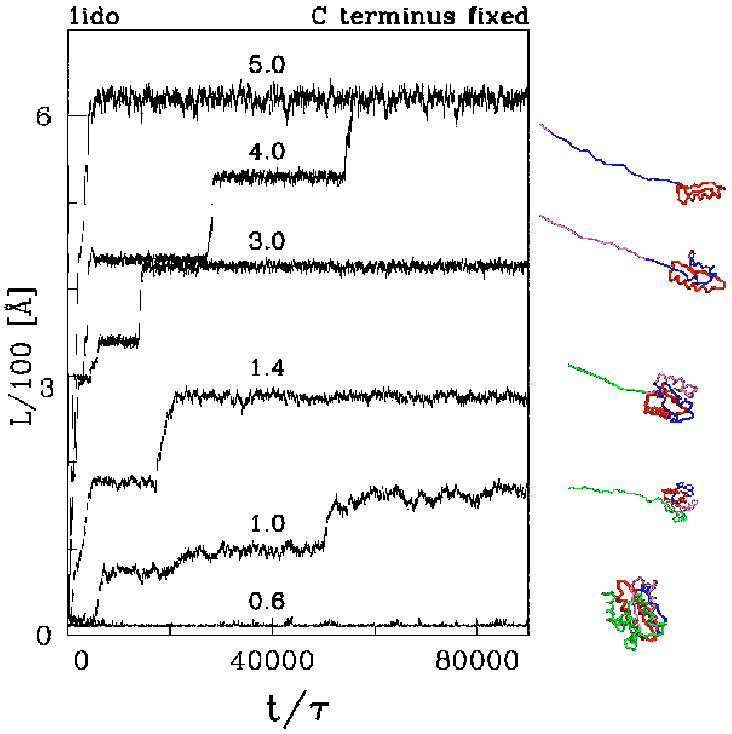
1ido
C terminus fixed
5.0
4.0
3.0
1.4
1.0
0.6
6
3
0
L/100 [Å]
0
40000
80000
t/τ

FIG. 3:

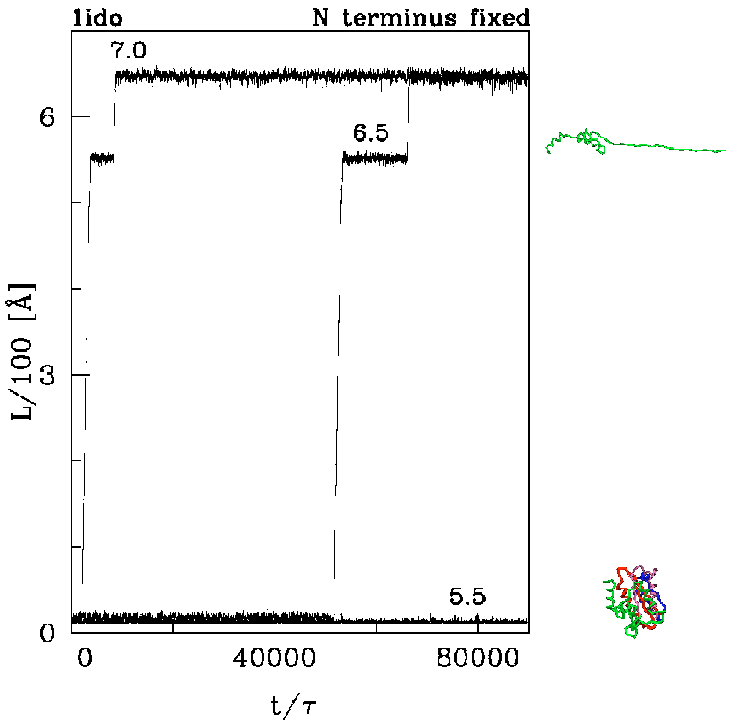

FIG. 4:

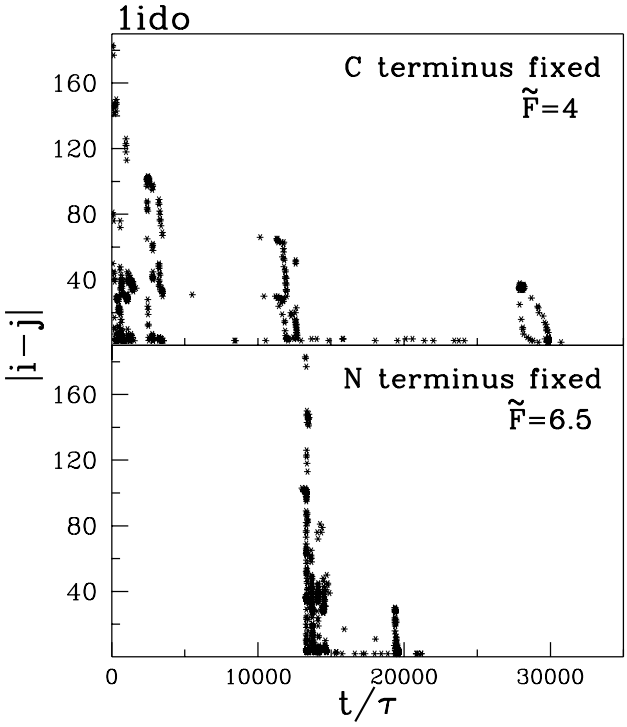
1ido
C terminus fixed
$\tilde{F}$=4
N terminus fixed
$\tilde{F}$=6.5
|i−j|
160
120
80
40
0
10000
20000
30000
t/τ

FIG. 5:

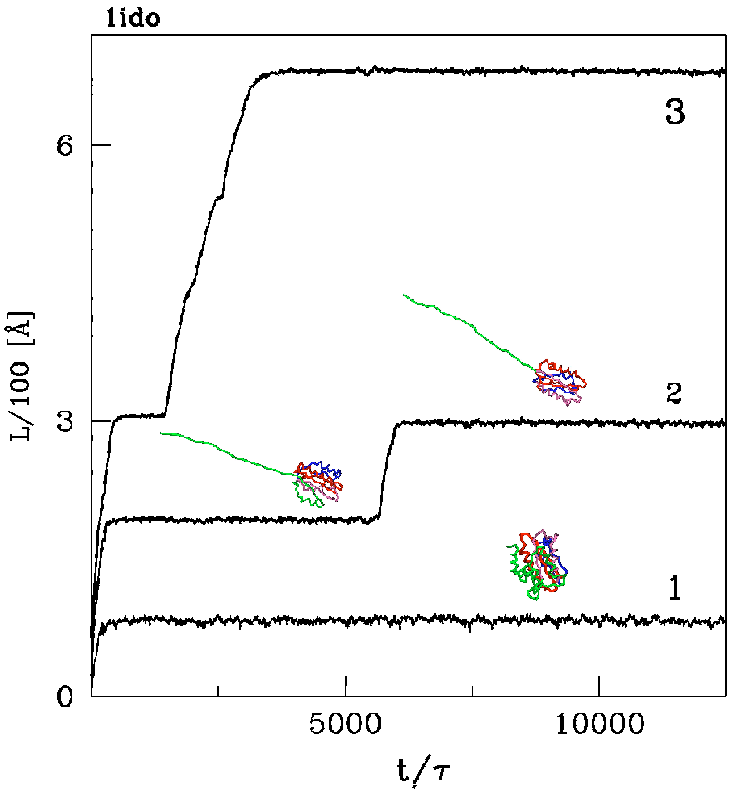

FIG. 6:

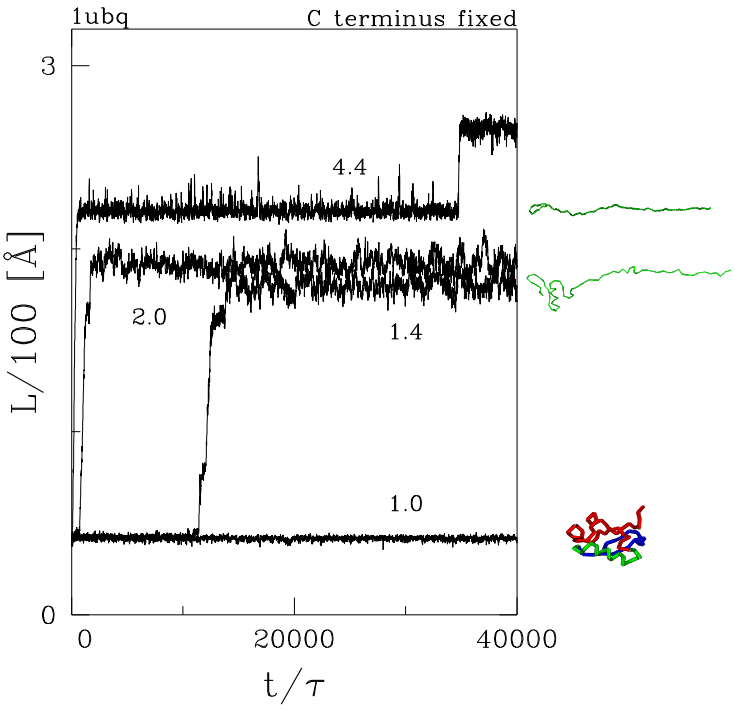

FIG. 7:

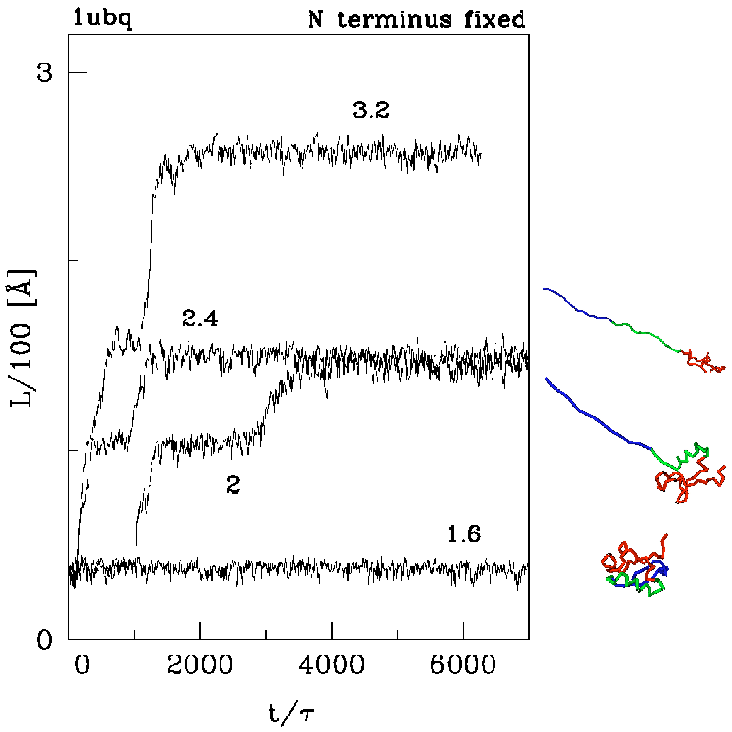

FIG. 8:

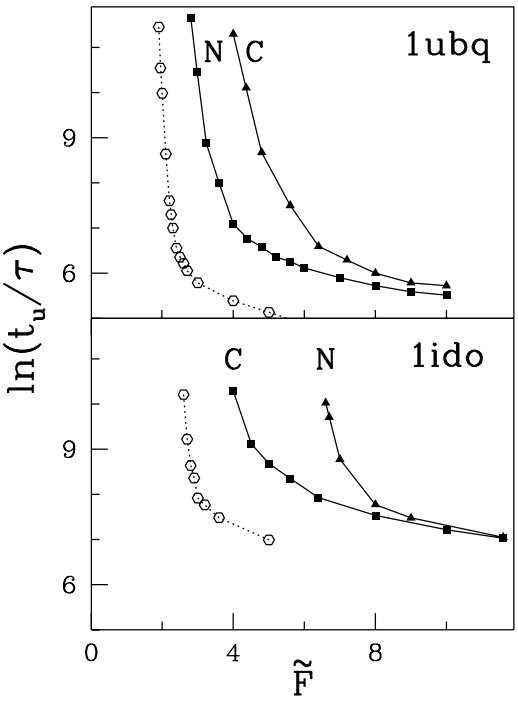

FIG. 9:

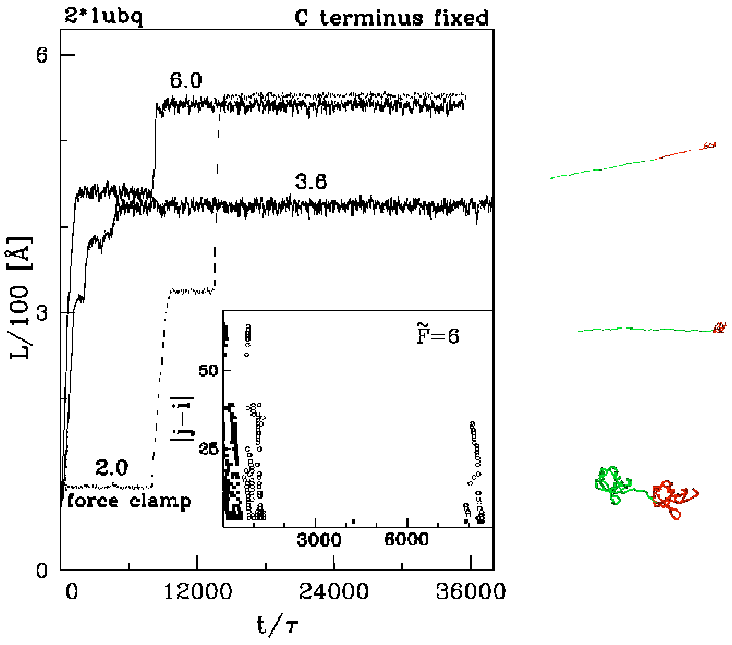
2*1ubq
C terminus fixed
6.0
3.6
2.0
force clamp
L/100 [Å]
t/τ
|j−i|
F̃=6
0
12000
24000
36000
3
6
25
50
3000
6000

FIG. 10:

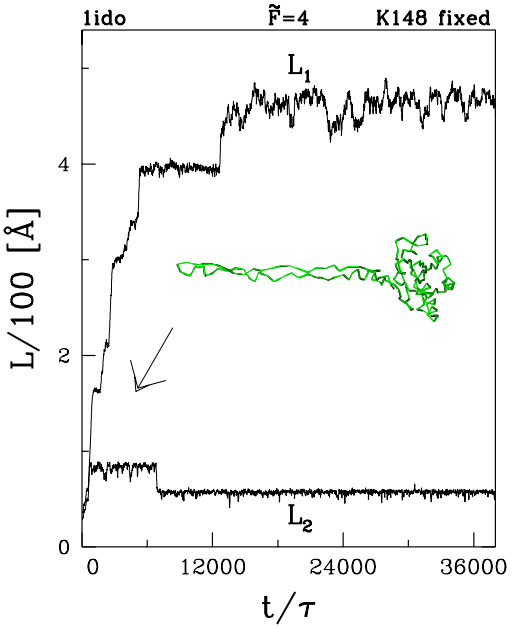
1ido
$\tilde{F}$=4
K148 fixed
L_1
L_2
L/100 [Å]
0
2
4
0
12000
24000
36000
t/τ

FIG. 11:

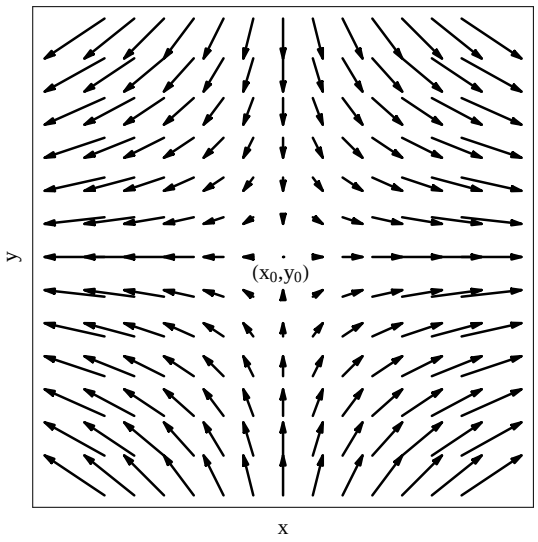

FIG. 12:

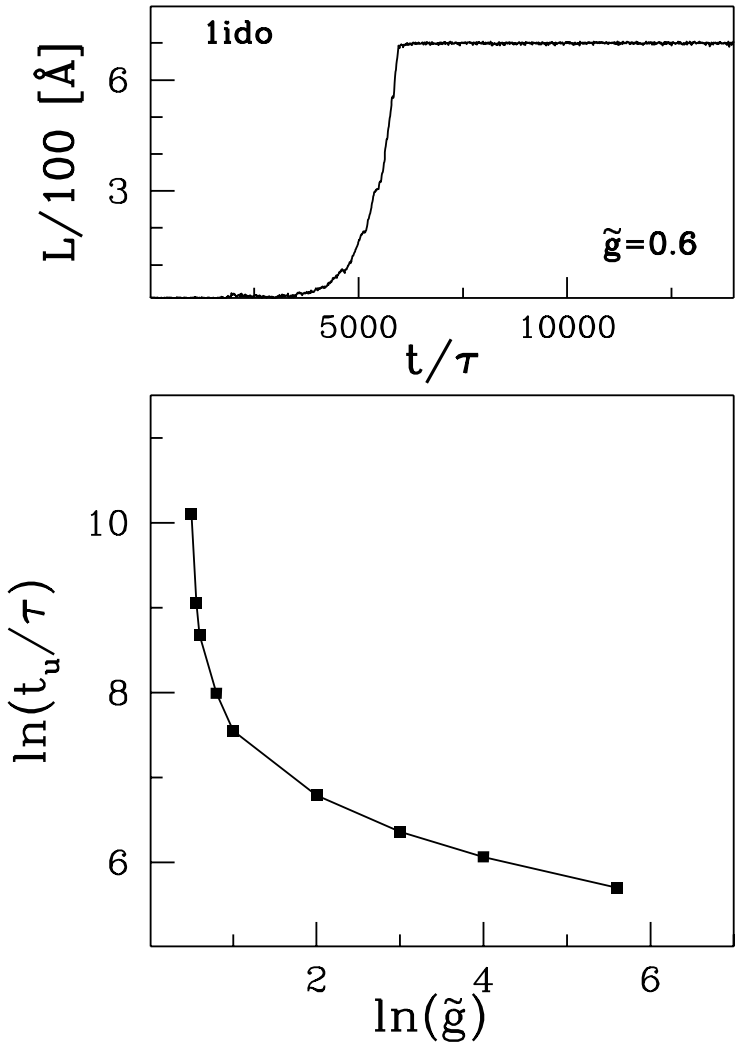
1ido
L/100 [Å]
6
3
$\tilde{g}$=0.6
5000
10000
t/τ
10
8
6
ln(t_u/τ)
2
4
6
ln($\tilde{g}$)

FIG. 13:

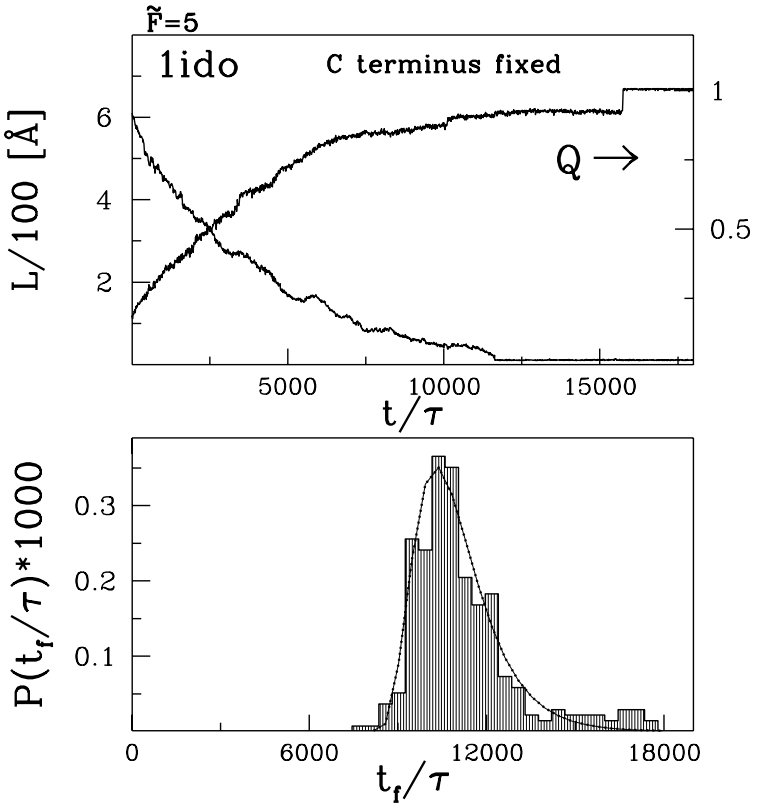
F̃=5
1ido
C terminus fixed
L/100 [Å]
6
4
2
1
0.5
Q →
5000
10000
15000
t/τ
P(t_f/τ)*1000
0.3
0.2
0.1
0
6000
12000
18000
t_f/τ

FIG. 14: